

A robust approach to IMRT optimization

Timothy C Y Chan¹, Thomas Bortfeld² and John N Tsitsiklis³

¹ Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

² Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

³ Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

E-mail: tcychan@mit.edu, tbortfeld@partners.org and jnt@mit.edu

Received 25 January 2006, in final form 26 January 2006

Published 4 May 2006

Online at stacks.iop.org/PMB/51/2567

Abstract

Managing uncertainty is a major challenge in radiation therapy treatment planning, including uncertainty induced by intrafraction motion, which is particularly important for tumours in the thorax and abdomen. Common methods to account for motion are to introduce a margin or to convolve the static dose distribution with a motion probability density function. Unlike previous work in this area, our development does not assume that the patient breathes according to a fixed distribution, nor is the patient required to breathe the same way throughout the treatment. Despite this generality, we create a robust optimization framework starting from the convolution method that is robust to fluctuations in breathing motion, yet spares healthy tissue better than a margin solution. We describe how to generate the data for our model using breathing motion data and we test our model on a computer phantom using data from real patients. In our numerical results, the robust solution delivers approximately 38% less dose to the healthy tissue than the margin solution, while providing the same level of protection against breathing uncertainty.

1. Introduction

The classical approach to dealing with uncertainty in intensity modulated radiation therapy (IMRT) has been to introduce a margin surrounding the tumour volume and incorporate it in the treatment planning. Such an approach, while producing a treatment that is robust against uncertainty, is over-conservative and necessarily increases the radiation exposure of healthy tissue and organs-at-risk. At the other end of the spectrum, if the treatment plan ignores uncertainty altogether, the result will typically be a non-uniform overall dose to the tumour with substantial cold spots. The work presented in this paper strikes a middle-ground between two conflicting objectives, ensuring the tumour receives the required dose, while exposing healthy tissue to less dose than the corresponding margin-induced treatment.

We will focus on intrafraction motion as our main source of uncertainty—breathing motion in lung tumours, to be precise—which has been observed in many studies (Langen and Jones 2001, Goitein 2004, Shimizu *et al* 2001, Shirato *et al* 2004, Seppenwoolde *et al* 2002). Accordingly, our concept of a margin will only pertain to intrafraction breathing motion uncertainty, ignoring other sources of uncertainty such as systematic errors, which need to be analysed separately. Our development is related to the idea of a ‘motion pdf’ (Trofimov *et al* 2005, Bortfeld *et al* 2002, Zhang *et al* 2004, Li and Xing 2000, Engelsman *et al* 2005), which uses a probability density function to define what proportion of time the tumour spends in each breathing phase (in reality, discretization leads to a probability mass function or pmf). Similar pmf models exist for interfraction motion (Unkelbach and Oelfke 2004). To use this pmf approach, it is necessary that ‘the motion ... is reproducible and stable during the treatment delivery’ (Trofimov *et al* 2005). In other words, using a pmf to account for tumour motion produces acceptable dose distributions provided the motion does not deviate significantly from what is expected. Our work here explores the consequences of violating this ‘negligible deviation’ assumption. In particular, we develop a framework to analyse uncertainty in the pmf that describes breathing motion, and show how to incorporate it into the inverse planning optimization to produce a treatment plan that is robust to these uncertainties. Throughout the paper, ‘robust’ will mean that no realizations of the uncertainty will cause a constraint (e.g., the minimum dose requirement in the tumour) to be violated. Our analysis focuses on the case where there is no uncertainty in the objective function. In the appendix, we discuss how to extend our framework to include uncertainty in the objective function.

In section 2, we introduce a model of data uncertainty which describes the possible variations in the breathing pmf to be protected against. In section 3, we describe the process of deriving a pmf from breathing motion data, and also show how to use multiple pmfs to produce bounds on the uncertainty model. In section 4, we present a mathematical formulation which represents the inverse planning problem at hand, with two distinctive features: first, it is a *robust* formulation, capable of mitigating the breathing uncertainty and producing a plan that satisfies our conflicting objectives of tumour coverage and healthy tissue sparing. Second, it is a *linear* formulation, ensuring that the robust problem is negligibly more complex than our original problem. In section 5, we illustrate results using our robust formulation and compare it to the margin approach, as well as to the nominal (no uncertainty) approach. Finally, in sections 6 and 7, we discuss some supplementary issues regarding our model and conclude with a future outlook of robust optimization in radiotherapy treatment planning.

2. Model of uncertainty

If motion is ignored during the inverse planning process, but is present during treatment, the resulting dose distribution will be averaged out (this is often referred to as ‘dose-blurring’ (Jiang *et al* 2003, Engelsman *et al* 2005, Goitein 2004, Bortfeld *et al* 2004)). In particular, the resulting distribution will be the result of convolving the static dose distribution with the motion pmf. To incorporate motion in the planning process, we can adjust the dose deposition matrix *before* we optimize (i.e., solve a deconvolution problem), so that the resulting static dose distribution convolved with the motion pmf matches the desired distribution. This process of adjusting for motion relies heavily on the fact that the pmf used in the planning will be the same as the one realized during treatment (and is known *a priori*), which has been the standard assumption with the convolution-based approach so far (Trofimov *et al* 2005, Zhang *et al* 2004, Unkelbach and Oelfke 2004, Li and Xing 2000, Engelsman *et al* 2005). However, as motion can vary from day to day and within treatment sessions, we need an approach to

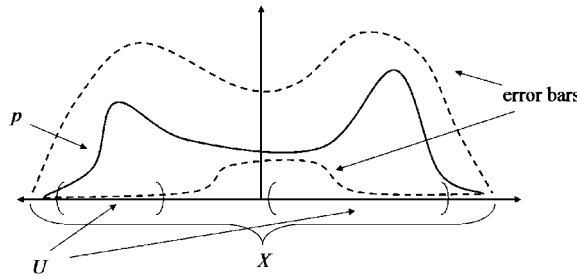


Figure 1. A visualization of the model of uncertainty.

mitigate the uncertainty of realizing different pmfs during treatment. Accordingly, the goal in this section is to introduce a model of uncertainty for the motion pmf. As opposed to assuming a fixed distribution of motion, we present a framework that is distribution-free, where we give bounds on the uncertainty, but make no assertion about the exact shape of the distribution. In a sense, the uncertainty in the problem has been moved from uncertainty *due* to motion to uncertainty *in* the motion.

The proposed uncertainty model consists of a ‘nominal’ pmf, p , surrounded by upper and lower ‘error bars,’ $p + \bar{p}$ and $p - \underline{p}$, respectively. For a loose interpretation, one can think of the nominal pmf as the mean of some random vector (random function in a continuous space) describing tumour motion while the error bars describe the extent of the variability of each component of this random vector. The assumption is that any realized pmf during treatment will be wholly contained within these error bars. Hence, \bar{p} and \underline{p} bound the deviations of the realized pmf above and below, respectively, the nominal pmf p . We define X (a finite set due to the discretized nature of the problem) to be the domain of the nominal pmf and error bars, and without loss of generality, we can take X to be just the subset of the domain where the pmf or the error bars are non-zero (i.e., the union of the supports of the pmf and error bars). We introduce an ‘uncertainty region’ $U \subseteq X$, which specifies the locations where the realized pmf is allowed to deviate from the nominal one. Thus, on the set $X \setminus U$, we assume that the realized pmf matches the nominal pmf. This model gives us the flexibility to address many scenarios. If a patient’s breathing is highly irregular, then the nominal data are less useful, so we could set $U = X$ and push the error bars farther apart. On the other hand, if a patient’s breathing is very regular, we could bring the error bars closer together, or, if the patient’s breathing is regular at exhale and irregular at inhale, then we could set U equal to the part of the domain representing inhale. Figure 1 illustrates the model.

Mathematically, we define the pmf uncertainty set P_U as

$$P_U = \left\{ \tilde{p} \in \mathfrak{N}^{|X|} : \tilde{p}(x) \in [p(x) - \underline{p}(x), p(x) + \bar{p}(x)], \forall x \in U; \right. \\ \left. \tilde{p}(x) \geq 0, \forall x \in X; \sum_{x \in X} \tilde{p}(x) = 1; \tilde{p}(x) = p(x), \forall x \in X \setminus U \right\}. \quad (1)$$

In what follows, we will derive a formulation of the IMRT optimization problem whose solution will remain feasible, for every realized $\tilde{p} \in P_U$. Before that, we discuss how to obtain the nominal pmf and error bars from tumour motion data.

3. From patient data to a model of uncertainty

In this section, we describe the method we use to create data suitable to be included in the optimization starting from tumour motion data. Note that the development of our framework is

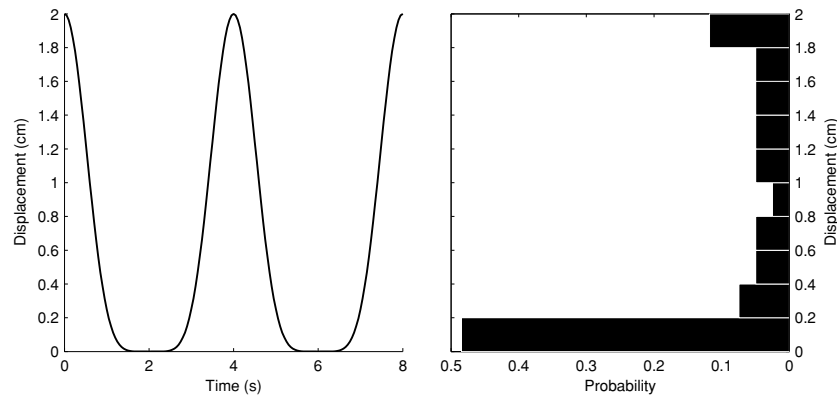


Figure 2. Generating a pmf from sinusoidal data.

independent of the data used, be it data collected with an external marker placed on the abdomen during breathing or internal tumour motion data obtained from a fluoroscopic marker. Ideally, in the optimization, a treatment planner would use data describing actual tumour motion induced by breathing, or external data plus the appropriate parameters to correlate internal motion to the external marker (Vedam *et al* 2003, Gierga *et al* 2005, Tsunashima *et al* 2004). Our analysis focuses on one-dimensional motion (e.g., along the superior–inferior direction), however, this approach can also be used as a basis for a similar analysis of motion in higher dimensions.

Given a trace which represents tumour motion (e.g., a sinusoidal curve), we can create a pmf by ‘horizontally aggregating’ the curve into a desired number of bins, that is, by simply constructing, on the vertical axis, a histogram of the data points along the trace (see figure 2). The generation of a pmf in this manner is quite standard (Lujan *et al* 1999, Trofimov *et al* 2005).

Next, we generate the error bars. Consider a family of pmfs, each created as described in the previous paragraph. Let F be this family of pmfs, denoted f_0, f_1, \dots, f_k , where f_0 is chosen to be the nominal pmf p . Overlay each of these pmfs on the same axes and consider the upper and lower envelopes created by these functions. Recall that $p + \bar{p}$ and $p - \underline{p}$ denote the upper and lower error bars, respectively. Thus, for each x where $f_i(x) > 0$ for some $i = 0, \dots, k$, (this is the set X) we define $\bar{p}(x)$ and $\underline{p}(x)$ so that

$$p(x) + \bar{p}(x) = \max_{i=0, \dots, k} f_i(x) \quad (2)$$

and

$$p(x) - \underline{p}(x) = \min_{i=0, \dots, k} f_i(x). \quad (3)$$

The intuition behind this construction is that we would like to be protected against any of the f_i from being realized during treatment and, in a sense, everything in between. So, $p(x) - \underline{p}(x) \leq f_i(x) \leq p(x) + \bar{p}(x)$ for all $x \in X$ and $i = 0, \dots, k$. Figure 3 illustrates the upper and lower error bars resulting from the displayed pmfs.

Let us summarize the procedure we have described in this section and ground it with examples of practical use. On a day prior to treatment, the planner will take reference 4D CT scans and simultaneously acquire tumour motion data from an internal or external marker (or both). Together, this information will be used to create a probability mass function to describe the motion of the tumour during breathing. Then, absent additional data from the patient, data from past patients will be used to create the error bars (more on this in section 6). The pmf

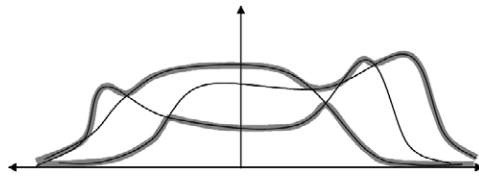


Figure 3. Generating the upper and lower error bars (bold lines) using multiple pmfs.

obtained from the patient will serve as the nominal pmf and, along with the error bars, will be fed as data into the optimization routine, which we describe next.

4. Robust formulation

In this section, we formulate the problem of IMRT optimization under uncertainty in the probability mass function describing tumour motion. We use a linear formulation (linear objective function and linear constraints—more examples in section 6) and some standard mathematical techniques to derive a tractable ‘robust counterpart’, so that a solution of this robust counterpart results in a dose distribution that is robust with respect to the uncertainty in the problem. That is, the resulting dose distribution satisfies the problem constraints under any realization of the uncertain pmf.

Linear programming is a powerful tool to solve optimization problems which arise in a multitude of applications. The many benefits of using a linear optimization framework include a guarantee of optimality at the termination of specialized algorithms, a beautiful theory of duality which gives insight into the sensitivity of solutions, and in our case, robust counterparts which are also linear and of only slightly larger problem size (thus, equally tractable). For a comprehensive introduction to the subject, consult any suitable text (Bertsimas and Tsitsiklis 1997).

First, let us consider the IMRT optimization problem, formulated in terms of beamlet optimization without motion. Let $D_{v,b}$ be the dose that voxel v receives per unit of intensity of beamlet b , w_b be the weight or intensity of beamlet b and θ_v be the desired or prescribed dose that voxel v should receive. Let \mathcal{V} be the set of all voxels under consideration (i.e., the entire phantom), \mathcal{T} be the set of voxels in the tumour and \mathcal{N} be the set of voxels of the normal (non-tumour) tissue. Let \mathcal{B} be the set of beamlets. We formulate our basic problem as follows:

$$\begin{aligned}
 & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} D_{v,b} w_b \\
 & \text{subject to} && \sum_{b \in \mathcal{B}} D_{v,b} w_b \geq \theta_v \quad \forall v \in \mathcal{T} \\
 & && w_b \geq 0 \quad \forall b \in \mathcal{B}.
 \end{aligned} \tag{4}$$

The objective is to minimize the total (or integral or mean) dose delivered to the phantom, while ensuring that all voxels in the tumour receive a specified level of dose. Typically, it is required that the tumour receives a uniform dose, so, in that case, $\theta_v = \theta$ for all $v \in \mathcal{T}$.

To account for motion using a pmf, we first fix a nominal position for each voxel and then consider voxel displacement relative to this nominal reference frame to create a matrix Δ , where $\Delta_{v,x,b}$ is the dose to voxel v , if it is shifted by an amount x from its nominal position, per unit intensity of beamlet b (note that $\Delta_{v,x,b} = D_{v,b}$ if $x = 0$). By defining a probability mass function $p(x)$, with domain X , that describes the fraction of time a voxel is displaced by

x relative to its nominal position, we can use it to calculate the motion-averaged dose to voxel v from beamlet b :

$$\sum_{x \in X} \Delta_{v,x,b} p(x).$$

We can reformulate the basic problem to include motion as follows:

$$\begin{aligned} & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \geq \theta_v \quad \forall v \in \mathcal{T} \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \quad (5)$$

We will call this the *nominal* problem. This formulation is essentially the same as the basic problem (4), except with a slightly modified (motion-compensated) D matrix. The above formulation assumes rigid-body motion, but we can generalize the problem to non-rigid-body motion by using a possibly different pmf, $p_v(x)$, for each voxel v (Li and Xing 2000).

Now, we introduce the model of uncertainty described in section 2. We would like our formulation to generate solutions that remain feasible under any allowable realization, within our model (1), of the uncertain pmf. That is, if the realized pmf lies in P_U (within the error bars $p + \bar{p}$ and $p - \underline{p}$, and only differs from the nominal pmf p on the uncertainty region U), then we wish to guarantee that every voxel in the tumour receives at least its prescribed dose. This leads us to the following *robust problem*:

$$\begin{aligned} & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \quad \forall \tilde{p} \in P_U \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \quad (6)$$

The objective function states that we are minimizing the integral dose to the phantom, assuming that the patient breathes according to p . Note that in the objective function we are ignoring the uncertainty in the pmf, which is fine if our objective is to minimize integral dose. This is discussed in further detail in section 6. So, in our problem, we focus on uncertainty in the constraints. That is, we would like to be robust against any uncertainty that might prevent voxels in the tumour from receiving at least their prescribed dose.

Formulation (6) involves infinitely many constraints (because the set P_U has infinitely many elements) but can be rewritten (Chan 2007) in the equivalent form

$$\begin{aligned} & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b + \beta_v(w) \geq \theta_v \quad \forall v \in \mathcal{T} \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}, \end{aligned} \quad (7)$$

where

$$\begin{aligned} \beta_v(w) = \min_{\hat{p}} &&& \sum_{b \in \mathcal{B}} \sum_{x \in U} \Delta_{v,x,b} \hat{p}(x) w_b \\ & \text{subject to} && \sum_{x \in U} \hat{p}(x) = 0 \\ & && -\underline{p}(x) \leq \hat{p}(x) \leq \bar{p}(x) \quad \forall x \in U. \end{aligned} \quad (8)$$

We now interpret the reformulated problem (7), (8). For each voxel v in the tumour, the constraint requires that the nominal dose plus $\beta_v(\mathbf{w})$ must be at least the prescribed dose. The quantity $\beta_v(\mathbf{w})$ as defined in (8) represents the additional (typically negative) dose realized by a worst-case variation, \hat{p} , from the nominal pmf under our model of uncertainty. To see this, fix some \mathbf{w} and suppose that the nominal pmf p is changed to a realized pmf $p + \hat{p}$. The objective function in problem (8) that defines $\beta_v(\mathbf{w})$ is the resulting change in the dose actually delivered to voxel v . The constraints in problem (8) are equivalent to the constraint $p + \hat{p} \in P_U$, and therefore simply describe the allowed variations \hat{p} . The optimization in (8) is then used to identify the worst possible change (reduction) in the delivered dose to voxel v .

For some additional intuition, it is instructive to identify the structure of the worst-case variation, \hat{p} , as far as a particular voxel v is concerned, given a vector \mathbf{w} of beamlet intensities. In the absence of the equality constraint in (8), the minimum would be attained by setting $\hat{p}(x)$ equal to $-p(x)$ for all x . However, due to the equality constraint, a negative value of $\hat{p}(x)$ at some x needs to be balanced by a positive value of $\hat{p}(x)$ at other choices of x . The objective function in (8) is minimized by setting $\hat{p}(x) = -p(x)$ at those x for which $\sum_{b \in \mathcal{B}} \Delta_{v,x,b} w_b$ is above a certain threshold (the ‘high’ dose locations). This observation motivates the following interpretation. Fix the static dose distribution generated by \mathbf{w} and consider voxel v as it moves to various locations within this distribution. If v is scheduled, according to the nominal p , to spend a certain amount of time in the high dose locations, then we need to protect against the situation where v spends less time there than nominally assumed, and this is precisely what is accomplished by our formulation. By protecting against the scenario where v spends more time in the low dose locations and less time in the high dose locations than the nominal pmf dictates, we are prohibiting v from receiving less than the desired dose θ_v .

Formulation (6), as written, is not a linear program due to the infinite number of constraints. However, using the theory of linear programming duality, we can construct the following equivalent, linear formulation:

$$\begin{aligned}
 & \underset{\mathbf{w}, \mathbf{q}, \mathbf{r}}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\
 & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b - \sum_{b \in \mathcal{B}} \sum_{x \in U} \Delta_{v,x,b} \underline{p}(x) w_b \\
 & && + \sum_{x \in U} \underline{p}(x) q_v - \sum_{x \in U} r_{v,x} \geq \theta_v && \forall v \in \mathcal{T} \\
 & && (\bar{p}(x) + \underline{p}(x)) q_v - r_{v,x} \leq \sum_{b \in \mathcal{B}} \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x)) w_b && \forall v \in \mathcal{T}, \forall x \in U \\
 & && q_v \text{ free} && \forall v \in \mathcal{T} \\
 & && r_{v,x} \geq 0 && \forall v \in \mathcal{T}, \forall x \in U \\
 & && w_b \geq 0 && \forall b \in \mathcal{B}.
 \end{aligned} \tag{9}$$

This is the *robust counterpart* of the robust problem (6). We have introduced some new variables and constraints, so indeed, the formulation has increased in size (relative to the nominal formulation). This increase is on the order of the number of voxels in the tumour multiplied by the number of voxels in the uncertainty region, but with a linear formulation and the current capability to solve linear programs, this will not significantly increase the computational overhead of solving this problem. We omit the proof of equivalence between (6) and (9) since it will be covered elsewhere (Chan 2007). Also, interested readers should consult other sources (Ben-Tal and Nemirovski 2000, 2002, Bertsimas and Sim 2004) for a more advanced introduction to robust optimization.

To summarize, we have used the model of data uncertainty described by the uncertainty region U and error bars $p + \bar{p}$ and $p - \underline{p}$ to formulate the robust IMRT optimization problem as a linear optimization problem. The optimal solution to (9) will minimize total dose delivered to \mathcal{V} , while ensuring that the tumour receives the desired dose, under all possible breathing pmfs within our uncertainty set. Practically speaking, we have a formulation that can be optimized by any standard linear programming solver and that produces a robust solution.

4.1. A continuum of robustness

For an intuitive, macroscopic understanding of our robust framework, consider the following. If there is no uncertainty in the problem, then we would just use the nominal formulation (5). However, with complete uncertainty (i.e., the amplitude of motion, or X , is bounded, but the pmf can be any function between 0 and 1 on this set), a margin appears to be the best choice for delivering a uniform dose to the tumour, since the corresponding static dose distribution is as uniform as possible. In fact, our robust formulation can produce *exactly* these two scenarios.

Using the robust formulation, with no uncertainty in the problem, we would set $U = \emptyset$ and optimize. The resulting static dose distribution would take advantage of the probability mass function being known with certainty to minimize dose delivered while satisfying the constraints. It is not difficult to see that by setting $U = \emptyset$, we have $\beta_v(\mathbf{w}) = 0$ and we recover the nominal problem (5). On the other hand, with complete uncertainty, we would want to make our uncertainty set, P_U , as forgiving as possible. That is, we would set $U = X$, $\bar{p}(x) = 1 - p(x)$ and $\underline{p}(x) = p(x)$ for all $x \in X$. In other words, the lower error bar would be 0 and the upper error bar would be 1 (in a continuous setting, the upper error bar would be infinitely high since it would need to bound all delta functions over the set X). We can prove mathematically (Chan 2007) that optimizing the robust formulation with these parameters produces a margin-like dose distribution—the same dose distribution produced by the solution of the following formulation:

$$\begin{aligned}
 & \underset{\mathbf{w}}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\
 & \text{subject to} && \sum_{b \in \mathcal{B}} D_{v,b} w_b \geq \theta_v && \forall v \in \mathcal{T} \\
 & && \sum_{b \in \mathcal{B}} D_{v,b} w_b \geq \mu_v && \forall v \in \mathcal{M} \\
 & && w_b \geq 0 && \forall b \in \mathcal{B}.
 \end{aligned} \tag{10}$$

This is the *margin* problem and \mathcal{M} represents the voxels in the margin. The proof of equivalence assumes the requirement of a uniform dose in the tumour and the margin (i.e., $\theta_v = \theta$ for all $v \in \mathcal{T}$, and $\mu_v = \theta$ for all $v \in \mathcal{M}$). Since we focus on ensuring the feasibility of a solution, we defined our margin to be the smallest set of voxels that, in addition to the nominal tumour voxels, would include the tumour under all feasible displacements allowed by our model of uncertainty. Formally, this can be written as

$$\mathcal{M} = \{v \in \mathcal{V} \setminus \mathcal{T} \mid v = \hat{v} + x \text{ for some } \hat{v} \in \mathcal{T}, x \in X\}. \tag{11}$$

Thus, by setting U , \bar{p} and \underline{p} appropriately, the robust framework can recover both the nominal and margin formulations, which are therefore special cases of our general robust formulation. Of course, with ‘intermediate’ values of U , \bar{p} and \underline{p} , the robust formulation will produce a solution that is neither strictly a nominal solution nor strictly a margin solution, but some combination of the two. This shows that our robust formulation is a generalization of current planning techniques. Furthermore, this implies that any information we can gather

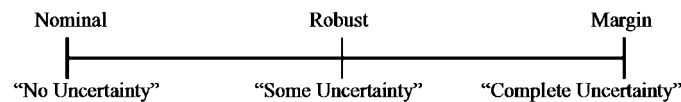


Figure 4. A visualization of the continuum of robustness.

about a patient to improve the error bars or uncertainty set from their most conservative values will result in a solution that is an improvement (measured via our objective function) over a margin.

We can visualize the flexibility of our robust framework as shown in figure 4. Imagine the nominal and margin problems as two extremes in a ‘continuum of robustness’, with the nominal representing ‘no uncertainty’, and the margin representing ‘complete uncertainty.’ A particular instance of the robust problem is, therefore, somewhere in the middle of this continuum and offers advantages of both extremes. It has the same capacity to mitigate uncertainty as the margin problem, but it also optimizes the resulting dose delivered to be as small as possible. Hence, a dose distribution produced using the robust formulation will be more robust to variations in the pmf than the nominal solution, but will also irradiate the normal tissue less than the margin solution. This is illustrated computationally in the next section.

5. Results

In this section, we motivate the use of robustness by considering what happens when the nominal formulation is used, first in the absence of uncertainty, and then with a realized pmf that is different from the planned one. Next, we show that while the margin solution is robust to the uncertainty in the realized pmf, it significantly increases the dose to the normal tissue. Finally, we show how the robust solution produces a dose distribution that, unlike the nominal solution, can combat uncertainty, while delivering less dose than the margin solution. Numerical results are shown at the end to supplement the illustrations.

For the illustrations below, we considered a one-dimensional phantom containing a 10 cm tumour surrounded by 10 cm of normal tissue on either side, and one beam orthogonal to the direction of motion. Voxels were 0.2 cm wide and beamlets were 0.5 cm wide, resulting in a total of 151 voxels and 28 beamlets. As shown in figure 5, the tumour motion is on the order of 2 cm peak to peak. The $D_{v,b}$ matrix was computed using error functions to model beamlet dose profiles (Engelsman *et al* 2005), and the mathematical formulations were optimized using ILOG’s CPLEX 9.0 solver (ILOG, Inc., Gentilly, France) through a Matlab interface (The MathWorks, Inc., Natick, Massachusetts). Each run took a few seconds on a 1.5 GHz computer with 1 GB of memory. While the geometry of this example is one dimensional, the illustrations below can be thought of as a slice of a two-dimensional phantom.

The pmfs and error bars were mined from actual patient data as described in section 3. In our analysis, we gathered data using Varian’s (Varian Medical Systems, Inc., Palo Alto, CA) real-time position management (RPM) system from four patients, totalling 95 traces. Each trace was baselined to its median value (so the resulting pmf has a median value of 0), and for simplicity, the average amplitudes of the RPM traces for the other three patients were scaled to approximately the average amplitude of the first patient. Although somewhat artificial, this was done to simulate the scenario where the population data are comparable in amplitude to the patient at hand. We used the planning RPM trace from the first patient to create the nominal pmf, the average of all 19 traces from the first patient to create the realized pmf and all 95 traces (one pmf per trace) from the four patients to create the error bars. To create the

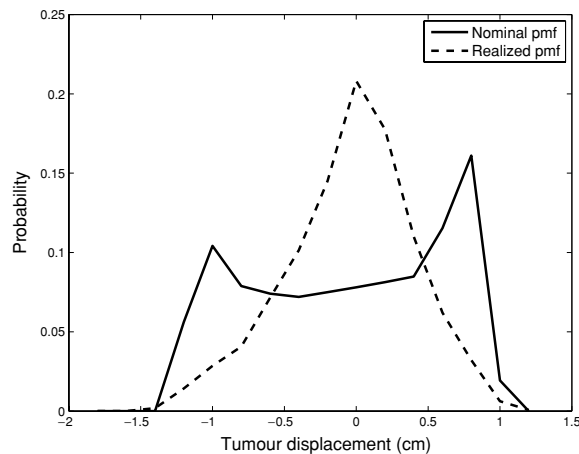


Figure 5. The pmfs used in the nominal, margin and robust formulation illustrations.

pmfs, we used a bin size equal to the width of a voxel because we focused on shifts of voxels by integer multiples of the voxel width. Since we used RPM data as a surrogate for tumour motion data, the amplitudes were scaled to be representative of large tumour motion.

The reader should keep in mind the following situation as we examine the figures below. If a patient breathes somewhat regularly, then we would expect to see the ‘bowl-shaped’ pmf in figure 5 (Lujan *et al* 1999). However, after many treatment days, the overall pmf may start to resemble a Gaussian distribution (George *et al* 2005, Engelsman *et al* 2005) due to differences in breathing patterns from day to day (‘realized pmf’ in figure 5). Thus, the principle behind applying the robust methodology in practice is as follows. If a patient breathes regularly during the reference session and we design the treatment plan using that pmf (nominal pmf), we will produce a plan with good tumour coverage and sharp dose gradients at the edge of the tumour, even if the cumulative pmf (realized pmf) over the course of the treatment is *significantly* different from the planned one. In the figures depicting the dose distributions of the three formulations, the pmf used in the planning for each is the nominal pmf in figure 5, while the pmf that is realized is either the nominal (part (a) of each figure) or the realized (part (b)) pmf.

In figure 6(a), we see the dose distribution from the nominal problem, using the same realized pmf as the nominal (planned) pmf. The tumour receives its prescribed dose in a largely homogeneous fashion, and there is a sharp falloff of the dose into the surrounding normal tissue. However, figure 6(b) shows severe hot and cold spots that result from the dose distribution if the realized pmf is different from the nominal one.

Figures 7(a) and (b) show the corresponding figures when the margin approach is used. As expected, the tumour receives a sufficient, homogeneous dose in both cases, but the tails of the dose distribution are much broader and extend further into the normal tissue, indicating a significant increase in the dose delivered there.

We now describe the results of the robust formulation. The uncertainty region was set equal to X , and the error bars shown in figure 8 depict the range of pmfs over which the robust solution is protecting. By comparing figures 9(a) and (b), we can see that the dose distribution from the robust plan is almost unchanged when a completely different pmf from the nominal pmf is realized. One interpretation of the robust dose distribution is that it is a combination of a margin-like distribution in the tumour to deal with uncertainty, and ‘horns’ at the edge

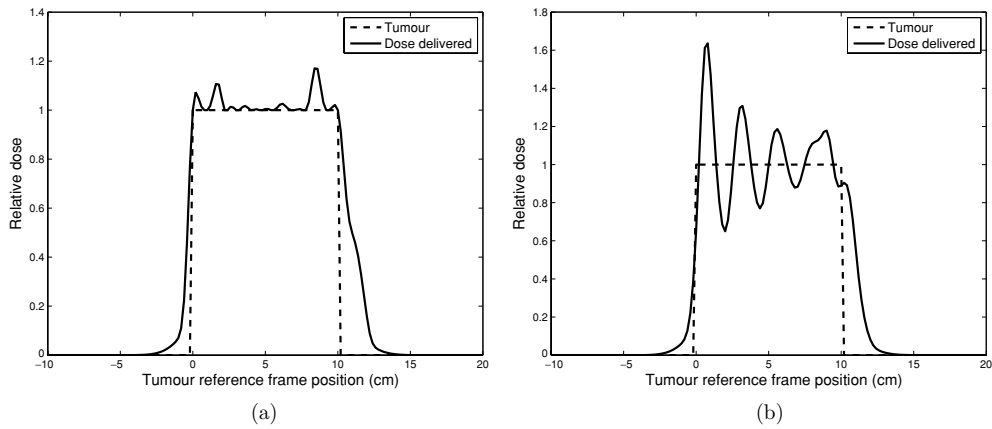


Figure 6. Dose distribution of nominal solution using the pmfs from figure 5. (a) Nominal pmf. (b) Realized pmf.

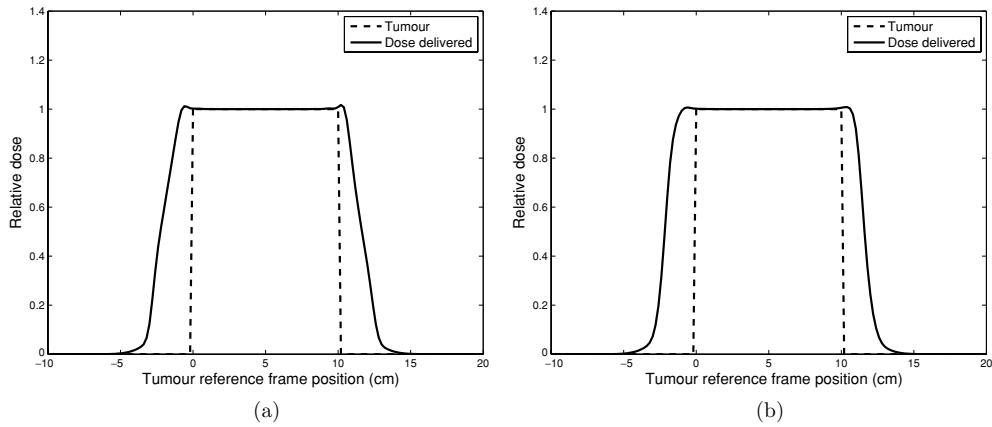


Figure 7. Dose distribution of margin solution using the pmfs from figure 5. (a) Nominal pmf. (b) Realized pmf.

of the tumour to produce a sharp dose gradient, combining the features of both the margin and nominal solutions. As you can see in figure 10, the intensity map (beamlet weights) corresponding to the robust solution is not overly complex.

Table 1 shows the numerical results corresponding to the previous figures. We can see that the robust solution delivers 8.55% less total dose and 38.03% less dose to the normal tissue than the corresponding margin solution. As expected, the nominal solution performs the best in terms of dose delivered, but the tradeoff is that it has no capability to mitigate uncertainty. There is a price to be paid to have protection from uncertainty, but the point is that by using the robust solution, we can pay a substantially lower premium than the cost of the margin solution, while achieving the same level of robustness to uncertainty.

6. Discussion

The mathematical framework introduced in this paper is based on the emerging field of robust optimization and, as is characteristic of this methodology, does not fix an underlying

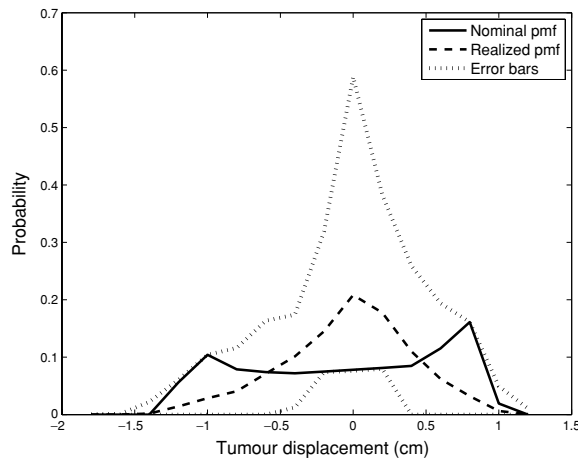


Figure 8. The pmfs and error bars used in the robust formulation illustrations.

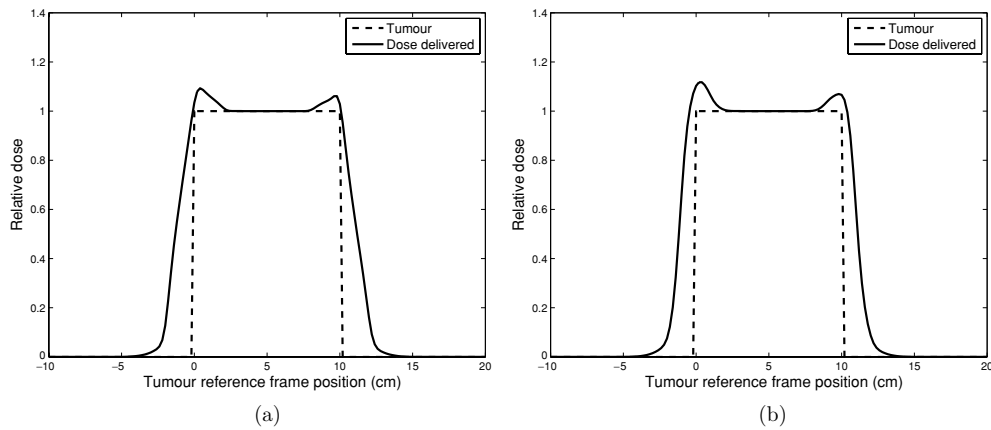


Figure 9. Dose distribution of robust solution using the pmfs and error bars from figure 8. (a) Nominal pmf. (b) Realized pmf.

Table 1. Comparing formulations on integral dose delivered to phantom and integral dose delivered to normal tissue under the realized pmf. The relative amount of dose delivered (%) by the various formulations is normalized to the margin formulation.

	Nominal	Robust	Margin
Dose to phantom	85.29	91.45	100.00
Dose to normal tissue	31.41	61.97	100.00

distribution for the uncertain quantities. Previous work typically assumed that the breathing motion during treatment was the same as during planning, and moreover, often assumed a Gaussian distribution for various uncertain parameters. While our analysis indicated that the cumulative breathing pmf may resemble a Gaussian after many fractions, we did not restrict our attention to just one distribution. Rather, the error bars shown in figure 8 indicate a range of distributions that are being protected against, including many distributions that are approximately Gaussian and, of course, many that are not. The distribution-free nature of our model is a key component of its flexibility.

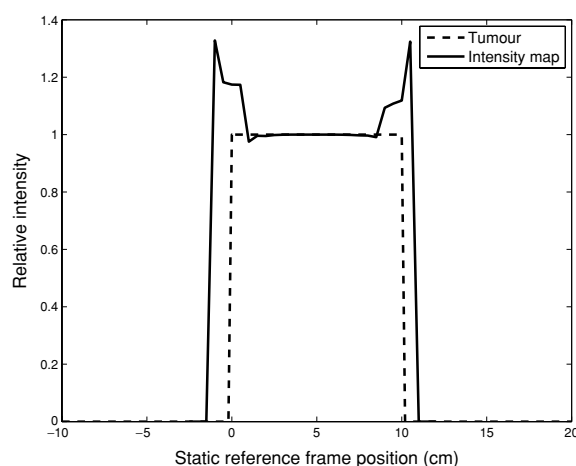


Figure 10. Intensity map corresponding to robust solution.

Directly incorporating robustness into treatment planning is an active topic of research (Baum *et al* 2006). Our work is similar in spirit to another recent study on robust IMRT (Chu *et al* 2005). The main difference is that our model of uncertainty is motivated by and constructed systematically from motion data. Another difference is that our approach ensures feasibility of the resulting solution under all realizations of uncertainty, as opposed to having feasibility with high probability. Mathematically, this translates to a formulation with lower complexity (a linear formulation) than the previous work (a second-order cone formulation). Our model uses the fact that the robust formulation remains a linear program to our advantage, to ensure that the computational overhead of solving the robust problem is marginally more expensive than the nominal problem. Because of their simplicity and ease of solution, linear formulations have been explored by many researchers (Bahr *et al* 1968, Hodes 1974, Sonderman and Abrahamson 1985, Langer and Leong 1987, Langer *et al* 1990, Morrill *et al* 1990, Rosen *et al* 1991, Shepard *et al* 1999, Holder 2003, Romeijn *et al* 2003).

A large variety of objective functions can be modelled using a linear framework. For parallel organs such as the lung (Niemierko 2006), it has been found that the mean dose is a suitable metric to consider (Theuws *et al* 1998). In this paper, we chose to minimize the total dose to the phantom instead of the normal tissue because our simple, one-dimensional phantom cannot represent all of the normal tissue surrounding the tumour (the tissue between the beam source and the tumour). In general, however, a simple modification allows us to minimize over \mathcal{N} instead of \mathcal{V} . While the standard quadratic objective cannot be cast in a linear framework, we can approximate it with the absolute value function or a piecewise linear objective (Romeijn *et al* 2003), both of which can be formulated in a linear fashion. Furthermore, minimizing the maximum dose to any voxel in a particular structure, or the weighted sum of any of the previous objectives can be modelled linearly. In fact, our current objective function can be considered a weighted sum of total dose to the tumour and total dose to the normal tissue. In the appendix, we show how to incorporate the objective of minimizing the maximum dose in our robust formulation, thereby extending our framework to deal with serial organs-at-risk such as the spinal cord (Niemierko 2006, Withers *et al* 1988, Wolbarst *et al* 1982). However, including biologically motivated quantities such as tumour control probability (TCP) or normal tissue complication probability (NTCP) (Niemierko 2006) in the optimization may necessitate a departure from a linear programming framework, and needs to be studied separately.

In addition to objective function modifications, other constraints can be incorporated into our formulation to model different requirements. We may include constraints on the complexity of the intensity map by constraining the difference between the minimum and maximum intensities, or by constraining the intensity difference between neighbouring beamlets. In a similar fashion, we can constrain the complexity or smoothness of realizable pmfs in our uncertainty set, which results in a less conservative robust solution. And of course, we can constrain the weighted sum of total/min/max dose delivered to various structures of interest. All of these requirements can be written linearly.

Regarding the data of the formulation, in principle, $\Delta_{v,x,b}$ could be calculated for each (v, x, b) triplet to most accurately represent the dose delivered under each possible voxel displacement scenario (McShan *et al* 2002, McShan and Fraass 2004). However this is unlikely to be practical given current computing capabilities. Instead, we can assume that we are working in a static coordinate system in which the dose distribution is fixed (for a given w), even though a given beamlet will pass through different densities of matter as structures move. It has been shown that such an approximation does not affect the dose calculation significantly (Bortfeld *et al* 2004, Engelsman *et al* 2001). Furthermore, we recognize that our problem is much smaller than typical clinical cases, but our intent is to illustrate the potential advantages of using a robust optimization approach and to be able to make effective comparisons with the commonly studied nominal and margin formulations.

While the pmf convolution method assumes that the dose is delivered over an infinite number of fractions, each witnessing an infinite number of breathing cycles, the error bars in our robust framework are general enough to partially mitigate this fractionation effect. Fractionation and a finite sampling of the breathing pmf implies that a pmf different from the nominal one should be realized in each treatment session. The error bars take into account this variation around the nominal pmf. This motivates the further study of our robust methodology in order to combat other types of uncertainty.

Regarding the nominal pmf used in the objective function, recall that we do not consider uncertainty in these data. Since our metric is integral dose, the particular pmf used in the objective becomes inconsequential as long as the phantom is large compared with the amplitude of motion. While different pmfs in the objective will result in slightly different computed integral doses, the difference will only be in the tails of the dose distributions (the dose outside of \mathcal{V}), which is small since the dose profiles of the beamlets fall off quickly (as error functions, in our case). A mathematically rigorous proof of this claim will be presented elsewhere (Chan 2007). For consistency across the three formulations, we used the same objective function in each, even though the margin formulation does not need to explicitly include a particular pmf. Typically, motion is implicitly included in the definition of the margin.

Our formulation is also compatible with other types of uncertainty such as setup error. For example, consider the tumour plus a setup margin (van Herk *et al* 2000) as the ‘entire’ tumour in the development presented in this paper. Then, our formulation ensures that the tumour plus margin receives the required dose under breathing motion uncertainty, as desired. Another way to include setup uncertainty is to modify the way the error bars are derived. By taking the pmfs used to create the error bars and shifting each one by amounts commensurate with the size of the setup margin around the tumour, we can create upper and lower error bars that protect against larger amplitudes of uncertain motion to compensate for the setup margin. While these types of extensions require no change in our framework to implement, a separate study will be needed to document the effects of including setup margins in our robust approach.

One possible practical challenge associated with using our methodology is the availability of suitable and reliable tumour motion data. In practice, a treatment planner is unlikely to

have a lot of data about a particular patient prior to treatment. In this case, the planner can use population data such as data from patients with similar breathing amplitudes, to construct the error bars. By including the nominal pmf of the patient at hand, generating the error bars according to section 3 guarantees that the nominal pmf is contained within the error bars. An alternative would be to compare the variations around the nominal pmf for past patients, and apply the relative deviations to each component of the current patient's nominal pmf to create the error bars. Also, it would be interesting to study whether or not we can create 'class solutions' for groups of individuals parameterized by, for example, similar breathing amplitudes. These topics deserve a separate analysis and will be areas of future study.

7. Conclusion

In this paper, we introduced a new robust methodology for dealing with IMRT optimization problems under uncertainty and considered the specific case of intrafractional uncertainty induced by breathing motion. We used the idea of a motion probability mass function along with an associated set describing the uncertainty of this pmf as our model of data uncertainty. A linear programming formulation of the basic IMRT optimization problem was presented and then systematically extended to the nominal, and finally, robust formulation. Our results showed that the robust solution was better suited to protect against uncertainty than the nominal solution, which did not account for uncertainty, while delivering less dose than the corresponding margin solution, which typically overcompensates for uncertainty. Thus, our robust framework produces what we term a 'continuum of robustness', allowing the user to vary his or her level of conservatism based on the patient at hand to produce solutions with features of both homogeneity and conformity. Overall, this work demonstrates the potential of using our robust optimization methodology in IMRT treatment planning to improve the sparing of healthy tissue while maintaining tumour coverage in the presence of uncertainty, and also the flexibility afforded to the treatment planner to make suitable decisions regarding trade-offs of conflicting objectives.

Future work will include using real tumour motion data in the optimization and applying the robust methodology to a clinical, 3D case. Possible extensions include optimization of proton therapy under uncertainty, considering the effects of fractionation on the delivered dose and realized pmfs, incorporating setup errors into the robust framework, and developing robust methodologies for gating and tracking.

Acknowledgments

The authors are grateful for insightful discussions with David Craft, Ben Martin and Alexei Trofimov. They would also like to thank Gregory Sharp for the help with acquiring and interpreting the RPM data.

This research was supported in part by the National Science Foundation under grant ECS-0312921, the Natural Sciences and Engineering Research Council of Canada, the United States–Israel Binational Science Foundation (BSF) under grant 2003275, and the National Cancer Institute under grant P01-CA21239 and grant R01-CA103904.

Appendix. Minimizing maximum dose as an objective

In this section, we show how to adapt the robust framework to serial organs, which may require (part of) the objective function to minimize the maximum dose delivered to voxels of

an organ-at-risk. We illustrate the method of constructing a linear formulation similar to (6) where the objective is now

$$\underset{w}{\text{minimize}} \quad \max_{\tilde{p} \in P_U, v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \quad (\text{A.1})$$

and P_U is defined as in (1). That is, the objective is to minimize the maximum dose that any voxel in the phantom receives, for all realizable \tilde{p} in our model of pmf uncertainty defined by P_U . Naturally, \mathcal{V} can be replaced by any other structure of interest. To recover a linear formulation, we first replace the objective function in (A.1) by the scalar z and add the constraints

$$z \geq \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \quad \forall v \in \mathcal{V}, \forall \tilde{p} \in P_U \quad (\text{A.2})$$

to the formulation. Then, a manipulation similar to the one in section 4 allows us to represent the problem as a linear program.

References

- Bahr G K, Kereiakes J G, Horwitz H, Finney R, Galvin J and Goode K 1968 The method of linear programming applied to radiation treatment planning *Radiology* **91** 686–93
- Baum C, Alber M, Birkner M and Nusslin F 2006 Robust treatment planning for intensity modulated radiotherapy of prostate cancer based on coverage probabilities *Radiother. Oncol.* **78** 27–35
- Ben-Tal A and Nemirovski A 2000 Robust solutions of linear programming problems contaminated with uncertain data *Math. Prog. A* **88** 411–24
- Ben-Tal A and Nemirovski A 2002 Robust optimization—methodology and applications *Math. Prog. B* **92** 453–80
- Bertsimas D and Sim M 2004 The price of robustness *Oper. Res.* **52** 35–53
- Bertsimas D and Tsitsiklis J N 1997 *Introduction to Linear Optimization* (Belmont: Athena Scientific)
- Bortfeld T, Jiang S B and Rietzel E 2004 Effects of motion on the total dose distribution *Semin. Radiat. Oncol.* **14** 41–51
- Bortfeld T, Jokivarsi K, Goitein M, Kung J and Jiang S B 2002 Effects of intra-fraction motion on IMRT dose delivery: statistical analysis and simulation *Phys. Med. Biol.* **47** 2203–20
- Chan T C Y 2007 *PhD Thesis* Massachusetts Institute of Technology (in preparation)
- Chu M, Zinchenko Y, Henderson S G and Sharpe M B 2005 Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty *Phys. Med. Biol.* **50** 5463–77
- Engelsman M, Damen E M F, Jaeger K D, van Ingen K M and Mijnheer B J 2001 The effect of breathing and set-up errors on the cumulative dose to a lung tumor *Radiother. Oncol.* **60** 95–105
- Engelsman M, Sharp G C, Bortfeld T, Onimaru R and Shirato H 2005 How much margin reduction is possible through gating or breath hold? *Phys. Med. Biol.* **50** 477–90
- George R, Keall P J, Kini V R, Vedam S S, Ramakrishnan V and Mohan R 2005 Is the diaphragm motion probability density function normally distributed? *Med. Phys.* **32** 396–404
- Gierga D P, Brewer J, Sharp G C, Betke M, Willett C G and Chen G T Y 2005 The correlation between internal and external markers for abdominal tumors: implications for respiratory gating *Int. J. Radiat. Oncol. Biol. Phys.* **61** 1551–8
- Goitein M 2004 Organ and tumor motion: an overview *Semin. Radiat. Oncol.* **14** 2–9
- Hodes L 1974 Semiautomatic optimization of external beam radiation treatment planning *Radiology* **110** 191–6
- Holder A 2003 Designing radiotherapy plans with elastic constraints and interior point methods *Health Care Manag. Sci.* **6** 5–16
- Jiang S B, Pope C, Jarrah K M A, Kung J H, Bortfeld T and Chen G T Y 2003 An experimental investigation on intra-fractional organ motion effects in lung IMRT treatments *Phys. Med. Biol.* **48** 1773–84
- Langen K M and Jones D T 2001 Organ motion and its management *Int. J. Radiat. Oncol. Biol. Phys.* **50** 265–78
- Langer M, Brown R, Urie M, Leong J, Stracher M and Shapiro J 1990 Large scale optimization of beam weights under dose-volume restrictions *Int. J. Radiat. Oncol. Biol. Phys.* **18** 887–93
- Langer M and Leong J 1987 Optimization of beam weights under dose-volume restrictions *Int. J. Radiat. Oncol. Biol. Phys.* **13** 1255–60
- Li J G and Xing L 2000 Inverse planning incorporating organ motion *Med. Phys.* **27** 1573–8

- Lujan A E, Larsen E W, Balter J M and Ten Haken R K 1999 A method for incorporating organ motion due to breathing into 3D dose calculations *Med. Phys.* **26** 715–20
- McShan D and Fraass B 2004 IMRT optimization with MIGA *Proc. XIVth Int. Conf. on the Use of Computers in Radiation Therapy* ed B Y Yi *et al* pp 232–5
- McShan D, Lynn K, Vineberg K and Fraass B 2002 Radiotherapy plan optimization accounting for set-up and motion uncertainty using a multiple instance geometry approximation (abstr.) *Med. Phys.* **29** 1257
- Morrill S, Rosen I, Lane R and Belli J 1990 The influence of dose constraint point placement on optimized radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **19** 129–41
- Niemierko A 2006 Biological optimization *Image-Guided IMRT* ed T Bortfeld *et al*, part II, chapter 5
- Romeijn H E, Ahuja R K, Dempsey J F, Kumar A and Li J G 2003 A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning *Phys. Med. Biol.* **48** 3521–42
- Rosen I, Lane R, Morrill S and Belli J 1991 Treatment plan optimization using linear programming *Med. Phys.* **18** 141–52
- Seppenwoolde Y, Shirato H, Kitamura K, Shimizu S, van Herk M, Lebesque J V and Miyasaka K 2002 Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **53** 822–34
- Shepard D M, Ferris M C, Olivera G H and Mackie T R 1999 Optimizing the delivery of radiation therapy to cancer patients *SIAM Rev.* **41** 721–44
- Shimizu S, Shirato H, Ogura S, Akita-Dosaka H, Kitamura K, Nishioka T, Kagei K, Nishimura M and Miyasaka K 2001 Detection of lung tumor movement in real-time tumor-tracking radiotherapy *Radiother. Oncol.* **51** 304–10
- Shirato H, Seppenwoolde Y, Kitamura K, Onimura R and Shimizu S 2004 Intrafractional tumor motion: lung and liver *Semin. Radiat. Oncol.* **14** 10–8
- Sonderman D and Abrahamson P G 1985 Radiotherapy treatment design using mathematical programming models *Oper. Res.* **33** 705–25
- Theuvs J C M, Kwa S L S, Wagenaar A C, Seppenwoolde Y, Boersma L J, Damen E M F, Muller S H, Baas P and Lebesque J V 1998 Prediction of overall pulmonary function loss in relation to the 3-D dose distribution for patients with breast cancer and malignant lymphoma *Radiother. Oncol.* **49** 233–43
- Trofimov A, Rietzel E, Lu H M, Martin B, Jiang S, Chen G T Y and Bortfeld T 2005 Temporo-spatial IMRT optimization: concepts, implementation and initial results *Phys. Med. Biol.* **50** 2779–98
- Tsunashima Y, Sakae T, Shioyama Y, Kagei K, Terunuma T, Nohtomi A and Akine Y 2004 Correlation between the respiratory waveform measured using a respiratory sensor and 3D tumor motion in gated radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **60** 951–8
- Unkelbach J and Oelfke U 2004 Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions *Phys. Med. Biol.* **49** 4005–29
- van Herk M, Remeijer P, Rasch C and Lebesque J V 2000 The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **47** 1121–35
- Vedam S S, Kini V R, Keall P J, Ramakrishnan V, Mostafavi H and Mohan R 2003 Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker *Med. Phys.* **30** 505–13
- Withers H R, Taylor J M G and Maciejewski B 1988 Treatment volume and tissue tolerance *Int. J. Radiat. Oncol. Biol. Phys.* **14** 751–9
- Wolbarst A B, Chin L M and Svensson G K 1982 Optimization of radiation therapy: integral-response of a model biological system *Int. J. Radiat. Oncol. Biol. Phys.* **8** 1761–9
- Zhang T, Jeraj R, Keller H, Lu W, Olivera G H, McNutt T, Mackie T and Paliwal B 2004 Treatment plan optimization incorporating respiratory motion *Med. Phys.* **31** 1576–86